Molecular BioSystems

Cite this: *Mol. BioSyst.*, 2011, **7**, 2890–2898

www.rsc.org/molecularbiosystems

PAPER

Characterization of efficient proteolysis by trypsin loaded macroporous silica†

Weichao Guo,[a] Hongyan Bi,[b] Liang Qiao,[b] Jingjing Wan,[a] Kun Qian,[a] Hubert H. Girault[b] and Baohong Liu*[a]

Received 8th April 2011, Accepted 11th June 2011
DOI: 10.1039/c1mb05140g

Tryptic digestion of proteins in trypsin loaded porous silica has been shown to be highly efficient. Enzymatic silica-reactors were prepared by immobilizing trypsin into macroporous ordered siliceous foam (MOSF) and into mesoporous SBA-15 silica which has a smaller pore size. The tryptic products from the silica reactors were analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS), and a higher proteolysis efficiency was obtained with MOSF. These results can be well interpreted by a sequential digestion model taking into account the confinement and concentration enrichment of both the substrates and enzymes within the silica pores. Proteins at low concentrations and proteins in urea and surfactant solutions were also successfully digested with the MOSF-based reactor and identified by MS. Considering that the immobilized trypsin could retain its enzymatic activity for weeks, this MOSF reactor provides many advantages compared to free enzyme proteolysis. As a proof-of-concept, the digest of a real complex sample extracted from the cytoplasm of mouse liver tissue using trypsin loaded MOSF yielded better results than the typical in-solution protocol.

Introduction

After years of development, proteomics has expanded its interface role to diverse research areas, such as infectious diseases, oncology, metabolic disorders, biomarker discovery, cardiovascular, pharma-toxicology and tissue imaging.[1–4] Mass spectrometry (MS) based protein identification is a key step of proteome research, often divided into bottom-up and top-down approaches.[5–7] Bottom-up proteomics involves proteolytic digestion of proteins prior to the analysis by MS, while top-down proteomics is a method using high performance MS and tandem MS to directly characterize the separated intact proteins.[5–7] Though top-down proteomics has drawn interest recently, bottom-up proteomics is still widely employed because of the instrumental and technical limitation of top-down strategies.

Proteolysis is a necessary step in bottom-up proteomics. The commonly used proteolysis methods include in-solution and in-gel protein digestion. These standard strategies based mainly on free enzymes work quite well in normal conditions, but perform poorly, for example, when digesting proteins at low concentrations,[8,9] and when digesting directly proteins in tissue[10] to keep the spatial information for MS imaging. Enzymes immobilized onto solid support are more stable and provide enhanced proteolysis efficiency.[11–13] Therefore, various strategies based on enzyme immobilization have been developed to overcome the shortcomings of in-solution or in-gel digestion, including proteolysis in microchip,[14,15] proteolysis in capillary[16] and *in situ* proteolysis on the sample plate of matrix assisted laser desorption ionization (MALDI) MS.[17] Recently, Bruening *et al.* designed a simple new method to create microporous reactors for fast protein digestion with sequential adsorption of poly(styrene sulfonate) and trypsin in nylon membranes.[18]

Meso- and macroporous materials possessing pore sizes from 2 nm to 50 nm and larger than 50 nm, respectively, have been systematically developed as catalysts to accelerate protein digestion and also to improve the biosensing response.[19,20] With theoretical predictions and experimental validations, it was demonstrated that macroporous materials such as macroporous ordered siliceous foam (MOSF)[21] can efficiently catalyze biochemical reactions by enriching reactants into its inner-pores and confining the reactions at the nano-scale.[19,22] It was proposed that the framework of MOSF with its large pore size and high pore volume could favor the adsorption of large biomolecules for example to facilitate enzyme–protein complex formation in macropores that may be hindered within smaller mesopores.[19,22]

[a] *Department of Chemistry and Institute of Biomedical Sciences, Fudan University, Shanghai, 200433, P. R. China. E-mail: bhliu@fudan.edu.cn; Fax: +86-21-65641740; Tel: +86-21-65642405*
[b] *Laboratoire d'Electrochimie Physique et Analytique, Station 6, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c1mb05140g

In this study, we demonstrate theoretically and experimentally that porous silica with a large pore diameter can catalyze proteolysis more efficiently than silica with smaller pores, and further propose an application of the MOSF-based proteolysis. Different from previous reports, where unmodified MOSF was directly added into a typical proteolysis solution as a catalyst,[19,22] herein trypsin is pre-loaded into the macropores of silica to form an enzymatic silica reactor. The immobilized enzymes are very stable and can be stored for weeks. During digestion, the enzyme-loaded porous materials were simply added into protein solutions to replace the free enzymes. After optimization of the loaded enzyme amount, a highly active enzymatic reactor was achieved that could be involved in the general proteolysis even in the presence of certain amounts of urea or surfactants. The trypsin-MOSF reactor was further applied to digest a real complex sample extracted from rat liver, and more proteins were identified from the silica reactor-assisted digestion than the typical in-solution digestion by high performance liquid chromatography electrospray ionization tandem MS (HPLC-ESI-MS-MS).

Experimental section

Reagents and chemicals

Trypsin, myoglobin and α-cyano-4-hydroxycinnamic acid (CHCA, 99%) were purchased from Sigma Chemical Co. (St. Louis, MO, USA). Acetonitrile (ACN, 99.9%) and trifluoroacetic acid (TFA, 99.8%) were purchased from Merck. Deionized water (18.2 MΩ cm) used for all experiments was obtained from a Milli-Q system (Millipore, Bedford, MA, USA). All other chemicals were of analytical grade and purchased from Shanghai Chemical Reagent Co. (Shanghai, China). All these reagents were used as received without further purification.

Preparation of porous silica and porous silica-immobilized trypsin, and the nanoreactor-based myoglobin digestion

MOSF and rod-like SBA-15 were synthesized as reported previously.[21,23] Scanning electron microscopy (SEM) micrographs were taken with a Philips XL30 microscope. Nitrogen sorption isotherms were obtained by using a Micromeritics Tristar 3000 analyzer at 77 K to calculate the specific surface area and total pore volume of the porous silica.

The porous material-trypsin reactor was prepared by mixing different amounts of trypsin (25–500 μg) with macroporous MOSF (1 mg) or mesoporous rod-like SBA-15 (1 mg) as a suspension in NH_4HCO_3 buffer (1 ml, 25 mM of NH_4HCO_3 in water, pH 7.8) under ambient conditions for 0.5 h. The suspension was centrifuged to remove excess trypsin, and then the particles were washed 2 times with the NH_4HCO_3 buffer to remove externally or weakly bound trypsin. The amount of trypsin immobilized was determined by UV absorption at 280 nm on a V-550 UV/Vis spectrophotometer by comparing the solution concentration before and after immobilization.

The silica reactor based myoglobin digestion was performed by incubating proteins (20 ng μl^{-1} or 2 ng μl^{-1}) that was dissolved in the NH_4HCO_3 buffer with the MOSF-trypsin or SBA-15-trypsin at the protease/substrate ratio of 1 : 3 (w/w) for 20 min at 37 °C. For comparison, the in-solution digestion with free enzymes was also performed for 20 min or overnight at 37 °C with an optimized enzyme/substrate ratio of 1 : 30 (w/w).[20]

Optimization and characterization of the enzymatic nanoreactors

To optimize the enzyme-loading amount, enzymatic silica reactors loaded with various amounts of trypsin ranging from 25 μg (mg MOSF)$^{-1}$ to 500 μg (mg MOSF)$^{-1}$ (maximum amount) were utilized to digest 20 ng μl^{-1} of myoglobin.

To evaluate how long the enzymatic activity of the silica reactor could be maintained, MOSF-trypsin reactors refrigerated for different periods, ranging from 1 day to 3 weeks was used for proteolysis of myoglobin.

To estimate the influence of surfactants on the digestion with the MOSF-trypsin reactor, myoglobin solutions in NH_4HCO_3 buffer with various concentrations of sodium dodecyl sulfate (SDS, 0.01%, 0.05% and 0.1% (w/v)) were employed as samples for digestion. Similarly, myoglobin solutions with different amount of urea (100 mM, 200 mM and 500 mM) were chosen as samples to evaluate the influence of urea on the nanoreactor proteolysis.

Preparation of the complex biological sample

Healthy rat liver tissue was obtained from the Liver Cancer Institute in Zhongshan Hospital, Fudan University. The sample was dissolved in a buffer containing urea (7 M) and sulfourea (2 M) mixture with protease inhibitors and phosphatase inhibitors (1mM PMSF, 0.2 mM Na_2VO_3, and 1 mM NaF). Then, the tissue sample was homogenized in an ice bath and vortexed for 30 min. The suspension was centrifuged at 18 000*g* for 1 h under 4 °C, and the supernatant was collected. The protein concentration of the extracted sample was determined according to the modified Bradford method described by Qu *et al.*[24] The extracted proteins were reduced with dithiothreitol (DTT) (20 mM) at 37 °C for 30 min and then alkylated with iodoacetamide (IAA) (25 mM) for another 30 min at room temperature in the dark.

2 mg of the extracted and denatured sample was taken for reverse phase liquid chromatography (RPLC) separation on a Shimadzu LC-20AD capillary pumping system by using a column of Agilent ZORBAX SB-C18 (4.6 × 250 mm, 5 μm, 300 Å, C18, Hypersil, EliteHPLC, China) for 120 min under 25 °C. 0.05% (v/v) TFA in water was used as mobile phase A and 0.05% (v/v) TFA in acetonitrile (ACN) was used as mobile phase B. All the collected fractions were lyophilized and stored in a refrigerator.

A fraction of the above mentioned separated rat liver samples (retention time from 77 min to 78 min, electronic supplementary information (ESI) Fig. SI-3†) was selected, lyophilized and redissolved in the NH_4HCO_3 buffer (25 mM, pH 7.8) with a final concentration of 100 μg ml^{-1}. For standard in-solution digestion, 100 μl sample proteins were incubated with trypsin with an enzyme/substrate ratio of 1 : 30 (w/w) for 3 h or 12 h under 37 °C. MOSF-based digestion was performed by incubating the same sample with trypsin-MOSF at an enzyme/substrate ratio of 1 : 3 (w/w) for 3 h under 37 °C.

MS analysis and database searching

The digestion products of myoglobin were analyzed on an Applied Biosystems 4700 proteomics analyzer (MALDI-TOF MS, Applied Biosystems, USA). Firstly, the sample solution (1 μl) was spotted onto the MALDI plate and dried under ambient conditions. Then, 0.5 μl of α-cyanohydroxycinnamic acid (CHCA) matrix solution (5 mg ml^{-1} CHCA, 0.3 mg ml^{-1} diammonium citrate in 0.1% TFA, 50% ACN, 49.9% H_2O solution) was added. Positive ion MALDI-TOF mass spectra were acquired. The instrument was operated at an accelerating voltage of 20 kV. A 200 Hz pulsed Nd:YAG laser at 355 nm was used. All mass spectra were obtained with an accumulation of 2000 laser shots under a laser intensity of 4000 instrument units and calibrated using an external calibration equation generated from the ion signal of tryptic digest of myoglobin. GPS Explorer software (Applied Biosystems) with Mascot (Matrix Sciences, London, UK) as a search engine and NCBI (version of 2009) as a database was used to analyze mass spectra. The peptide mass tolerance was set to 200 ppm. The signal to noise threshold was set as 3. Molecular weight search (MOWSE) score was calculated by the Mascot to show the protein identification accuracy.

For the rat liver fraction, the digested peptides were analyzed by RPLC-ESI-MS-MS using an LTQ mass spectrometer (Thermo-Fisher, CA). For database searching, the tolerance was set as ±0.6 Da for MS and ±0.8 Da for MS-MS; maximum one miss cleavage site of tryptic proteolysis was specified; and the charge states for ionized peptides were assumed as 2, 3 and 4. Peptides matched with significant homology ($p < 0.05$) were considered as identified peptides. Proteins were identified based on at least two significant peptides.

Results and discussion

Proteolysis with free enzyme and immobilized enzyme

Enzymes can be immobilized on porous silica due to non-covalent adsorption, such as electrostatic forces and hydrophobic–hydrophilic interactions. Compared to proteolysis induced by enzymes immobilized on beads or monoliths, proteolysis induced by enzymes immobilized in nanopores shows high efficiency because of the nano-confinement effect, where the fragments of the proteolysis are confined within a nanovolume and are therefore more likely to be further digested.[19] If there is specific affinity between the substrate and the protein, proteolysis could be more efficient due to the quick mass transport process.[19] Here, two different materials, macroporous ordered siliceous foams (MOSF) and mesoporous rod-like SBA-15, were employed as matrices to entrap enzymes, respectively. The MOSF possesses a three dimensional structure of uniform hexagonal arrayed columns with a pore diameter of about 100 nm, while SBA-15 has a much smaller pore size of only ~10 nm. Scanning electron microscope images and the physical parameters of these materials are shown in Fig. SI-1 and Table SI-1 in the ESI.† 100 μg of trypsin was immobilized onto 1 mg of each porous material, separately, to form the enzymatic silica reactor.

Myoglobin, known as a protein resistant to tryptic digestion,[25] was selected as the model substrate, and denatured at 100 °C for 5 min prior to digestion. The peptides mass fingerprinting (PMF) results of in-solution and in-reactor digestion are compared in Fig. 1. After being incubated at 37 °C for 20 min under shaking, 16 peptides were successfully identified with high signal-to-noise (S/N) ratio in the PMF spectrum from the MOSF-based proteolysis of 20 μg ml^{-1} myoglobin (Fig. 1a). In contrast, only 4 peptides were generated and identified by MS with strong noises using the in-solution digestion strategy under the same experimental conditions (Fig. 1b), and 11 peptides were produced and identified by MS with good S/N ratio using the typical overnight (12 h) in-solution digestion protocol (Fig. 1c). The SBA-15-assisted proteolysis showed results slightly better than that of the in-solution digestion but not as good as that of the MOSF-catalyzed proteolysis, where 9 peptides were identified after a 20 min digestion (Fig. 1d).

From these results, it is clear that both MOSF and SBA-15-immobilized trypsin can be used for fast proteolysis, while the MOSF-trypsin induces more efficient tryptic digestion than SBA-15-trypsin. This phenomenon can be interpreted by using the modified sequential digestion mechanism proposed recently.[19] According to our previous publication, two factors can influence the proteolysis efficiency of the in-nanopore digestion: the effectiveness factor η and the $V_{bulk}/V_{material}$ ratio, where V_{bulk} is the bulk solution volume of the digestion system and $V_{material}$ is the total pore volume of the nanoporous materials employed to immobilize enzyme.[19] For MOSF and SBA-15, the reported $V_{material}$ are quite alike: V_{MOSF} = 1.3 cm^3 g^{-1} and $V_{SBA\text{-}15}$ = 1.14 cm^3 g^{-1}, Table SI-1†, indicating a similar $V_{bulk}/V_{material}$ ratio for the proteolysis catalyzed by the same amount of MOSF or SBA-15. Therefore, the effectiveness factor η would mainly regulate the final proteolysis efficiencies of both reaction systems. The larger the effectiveness factors, the faster the proteolysis. η is expressed as:[19]

$$\eta = \frac{k_{in}/k_{out}}{1 + \frac{k_1 k_2}{k_{-1}+k_2} \times \frac{(1+k_{in}\delta/D)E_{local}}{k_{out}}} \quad (1)$$

where k_{in} and k_{out} are the ingress and egress rate constants (m s^{-1}), δ is the thickness of the diffusion layer around the porous materials, D is the protein diffusion coefficient in water (m^2 s^{-1}), E_{local} is the local concentration of enzymes in the nanopores of the porous material, and k_1, k_{-1}, k_2 are enzymatic reaction rates as explained in eqn (2).

$$\begin{gathered} S_{bulk} \xrightarrow{\text{diffusion}} S_{surface} \underset{k_{out}}{\overset{k_{in}}{\rightleftharpoons}} S_{local} \xrightarrow{\text{digestion}} P \\ E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_2} E + P_1 + P_2 \end{gathered} \quad (2)$$

where S is the substrate protein for digestion, E is the enzyme, P is digested peptides and C is the intermediate complex formed between enzyme and protein.

Eqn (1) can be further simplified as:

$$\eta = \frac{K_{ads}}{1 + \frac{(1+D_a)E_{local}k_2}{k_{out}K_M}} \quad K_M = \frac{k_{-1}+k_2}{k_1} \quad (3)$$

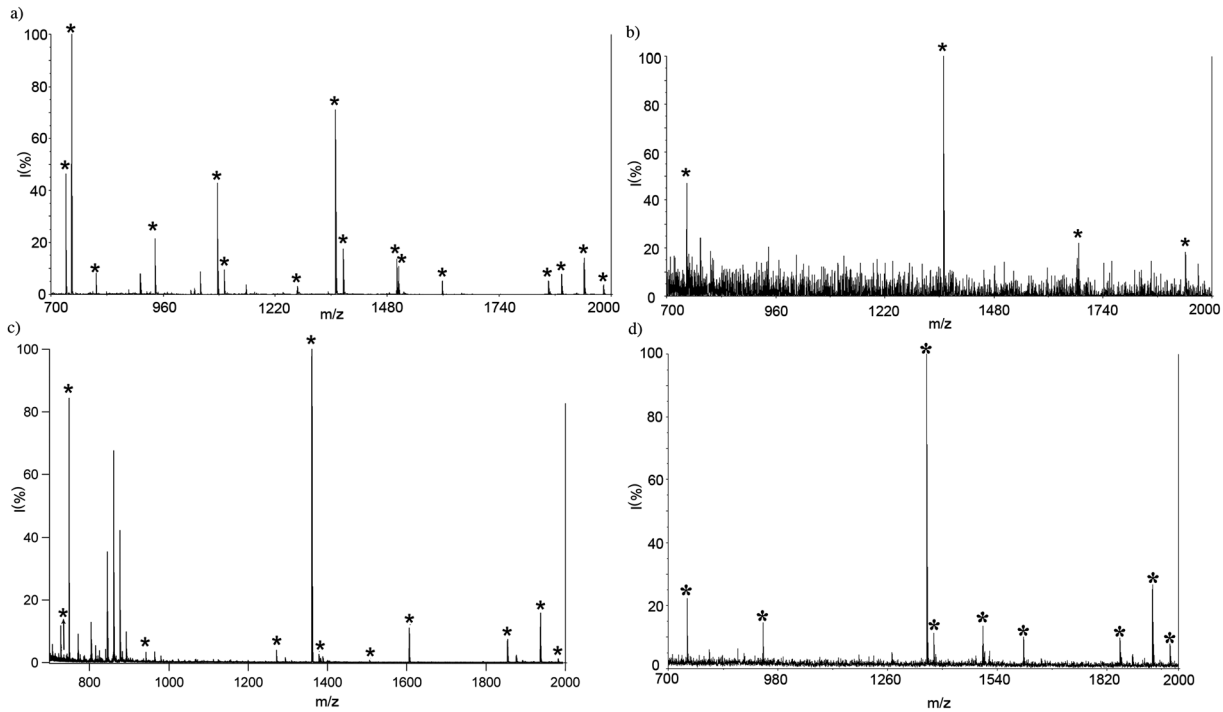

Fig. 1 PMF spectra of peptides obtained from (a) MOSF-assisted tryptic digestion products of myoglobin after 20 min incubation, (b) the in-solution tryptic digestion products of myoglobin after 20 min and (c) 12 h incubation, and (d) SBA-15-assisted tryptic digestion products of myoglobin after 20 min incubation. The peaks marked with * represent identified peptides of myoglobin. The concentrations of myoglobin used above are all 20 ng μL^{-1}.

where K_M is the Michaelis–Menten constant, K_{ads} is the adsorption constant (k_{in}/k_{out}), and D_a is the Damköhler number defined by:

$$D_a = \frac{k_{in}\delta}{D} \tag{4}$$

Comparing the proteolysis with MOSF and SBA-15-immobilized enzymes, K_M, k_2 and D should be the same in both reaction systems, because they are constants depending only on the nature of the enzyme, protein and solution employed. The thickness of the diffusion layer, δ, can be considered in a first approximation as 1 μm, which is comparable with the particle sizes of both MOSF and SBA-15. Since the enzyme loading amount is always 100 μg (mg material)$^{-1}$, and the $V_{material}$ is similar for MOSF and SBA-15, E_{local} in both reaction systems are also similar. In the present experimental conditions, the silica surface should be negatively charged in a buffer at pH 8,[22] therefore the counter positive ions would be attracted close to the silica surface to form a Gouy–Chapman electric double layer with a thickness of ~10 nm (details in ESI SI-2†). Considering that the pore diameter of SBA-15 is only ~10 nm comparable with the electric double layer thickness, and that the myoglobin with a molecule size of ~3.5 nm[26] in diameter should also be surface charged, we can assume that the ingress rate is limiting and $D_a \ll 1$ for SBA-15. In contrast, as the pore diameter of MOSF is ~100 nm, which is much bigger than both the electric double layer thickness and the protein size, we can assume that the diffusion rate is limiting and $D_a \gg 1$ for MOSF. The effectiveness factor η of MOSF catalyzed proteolysis can then be simplified as:

$$D_a^{MOSF} \gg 1$$

$$\eta_{MOSF} = \frac{K_{ads}}{1+\frac{(1+D_a^{MOSF})E_{local}k_2}{k_{out}^{MOSF}K_M}} = \frac{K_{ads}}{1+\frac{D_a^{MOSF}E_{local}k_2}{k_{out}^{MOSF}K_M}} = \frac{1}{1/k_{ads}+\frac{\delta}{D}K_E}$$

$$K_E = \frac{E_{local}k_2}{K_M} \tag{5}$$

where K_E can be named as the enzymatic factor that only depends on the nature of used enzyme and protein. Similarly, η of rod-like SBA-15 can be written as:

$$D_a^{SBA\text{-}15} \ll 1$$

$$\eta_{SBA\text{-}15} = \frac{K_{ads}}{1+\frac{(1+D_a^{SBA\text{-}15})E_{local}k_2}{k_{out}^{SBA\text{-}15}K_M}} = \frac{K_{ads}}{1+\frac{E_{local}k_2}{k_{out}^{SBA\text{-}15}K_M}} = \frac{1}{1/K_{ads}+\frac{1}{k_{in}^{SBA\text{-}15}}K_E} \tag{6}$$

Considering that MOSF and SBA-15 are both porous silica, the surface chemical and electronic properties of the two materials should be the same. Assuming that the enzymes and proteins are always adsorbed on or close to the pore inner walls, a similar K_{ads} value can be predicted for the reaction systems with both MOSF and SBA-15. Because $D_a^{SBA\text{-}15}$ is much smaller than unit, $1/k_{in}^{SBA\text{-}15}$ is much bigger than δ/D, and thereby $\eta_{SBA\text{-}15}$ is smaller than η_{MOSF}. With a bigger value of effectiveness factor, the MOSF immobilized trypsin can

then induce more efficient proteolysis than SBA-15 immobilized trypsin, consistent with the experimental results in Fig. 1.

Optimization of trypsin-loading amount on MOSF

To further apply the MOSF-immobilized trypsin reactor in proteolysis, the trypsin-loading amount on MOSF was optimized. It has been found that the maximum loading-amount for trypsin is 26 μmol (g MOSF)$^{-1}$, and that ~95% of the maximum adsorption can be achieved in less than 1 min.[22] However, it is not true that the high amount of trypsin immobilized on MOSF would result in satisfactory digestion. Too many trypsin molecules adsorbed into the nanopores may obstruct the adsorption of proteins, leading to poor proteolysis. Based on the hypothesis, five enzyme-loading amounts were tested, ranging from 25 μg (mg MOSF)$^{-1}$ to 500 μg (mg MOSF)$^{-1}$ (maximum loading amount). The results demonstrate that the MOSF-trypsin reactor with an enzyme-loading amount of 100 μg (mg MOSF)$^{-1}$ has the highest proteolytic activity, as shown in Table 1 and Fig. 2. According to the results, a trypsin loading-amount of 100 μg (mg MOSF)$^{-1}$ was selected for the subsequent experiments.

Table 1 Comparison of the digestion of 20 ng μL^{-1} myoglobin induced by MOSF-trypsin reactors with different enzyme loading-amounts

Trypsin amount/ μg (mg MOSF)$^{-1}$	MOWSE score[a]	Amino acid sequence coverage (%)	Peptides matched
500	169	64	11
200	196	74	12
100[b]	225	86	16
50	128	57	8
25	F[c]	8	2

[a] The molecule weight search score was calculated by the Mascot search engine. The higher the score, the higher the confidence level that the correct protein has been identified. [b] The mass spectrum for trypsin loading amount of 100 μg (mg MOSF)$^{-1}$ is shown in Fig. 1a. [c] F means that the identification failed, where Mascot did not give a significant analysis result.

Efficient protein digestion catalyzed by MOSF-trypsin

With the optimization of enzyme loading amount, the MOSF-trypsin reactor was used to induce efficient proteolysis of proteins at low concentrations. As shown in Fig. 3, the digestion result obtained after proteolysis for 20 min with MOSF-trypsin is much better than that obtained by using conventional in-solution digestion (20 min), indicating that the silica reactor can be used for the analysis of low-abundant proteins for proteome research. The mass spectrum of proteolysis products from 20 min digestion of myoglobin (2 ng μl^{-1}) catalyzed by MOSF-trypsin shows confident identification of 8 peptides with a MOWSE score of 117 and protein sequence coverage of 53%, according to the mass spectrum Mascot analysis result. In contrast, the in-solution digestion of 2 ng μl^{-1} myoglobin yields only three peptides identified by MS with very low S/N ratios after 20 min incubation, and the corresponding mass spectrum could not be analyzed by Mascot.

Stability of the MOSF-trypsin reactor

To investigate the stability of trypsin adsorbed into MOSF macropores, digestion of myoglobin (20 ng μL^{-1}) with the MOSF-trypsin reactors refrigerated for certain days was performed. Different time intervals, ranging from 1 day to 3 weeks, were tested as shown in Table 2. Fig. 4a–c show mass spectra of peptides digested from myoglobin with the catalysis of MOSF-trypsin after refrigerated storage for 1, 7, and 21 days. The MOWSE scores obtained by Mascot from various mass spectra are plotted against the storage time to

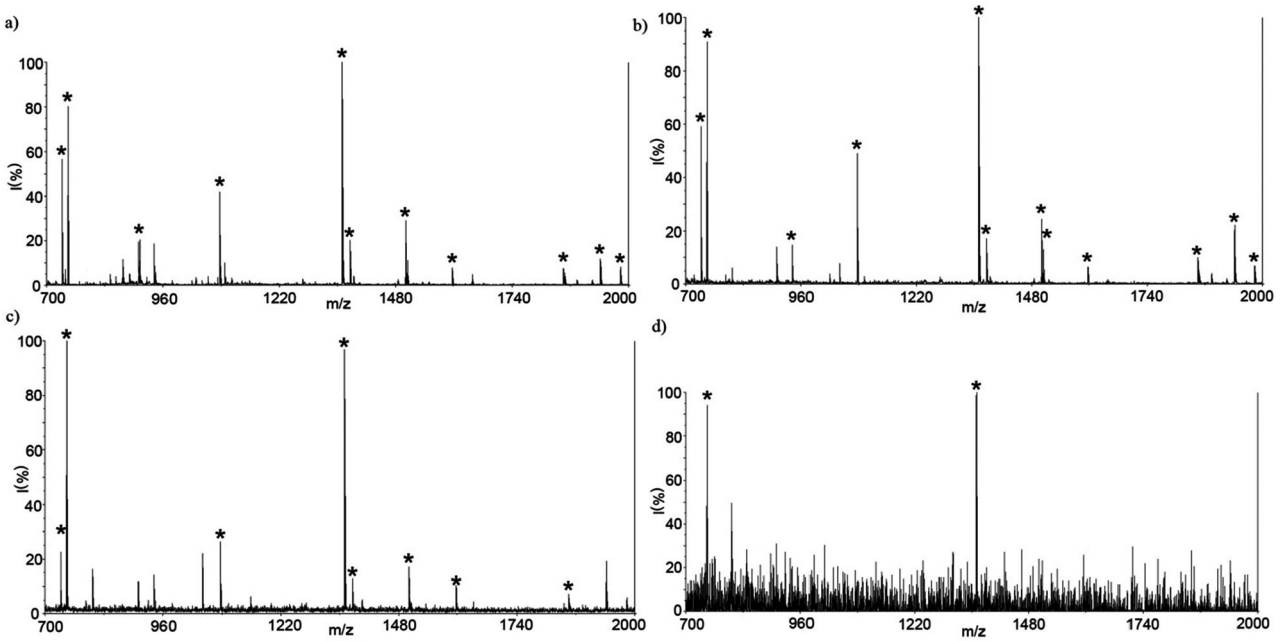

Fig. 2 Mass spectra of proteolysis products from 20 min digestion of myoglobin (20 μg ml^{-1}) catalyzed by the MOSF-trypsin reactors with enzyme loading amounts of (a) 500 μg (mg MOSF)$^{-1}$, (b) 200 μg (mg MOSF)$^{-1}$, (c) 50 μg (mg MOSF)$^{-1}$, and (d) 25 μg (mg MOSF)$^{-1}$.

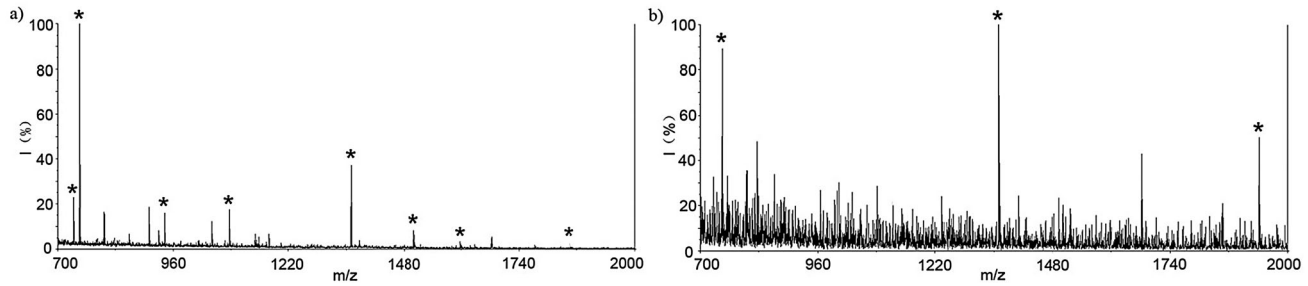

Fig. 3 Mass spectra of proteolysis products from (a) 20 min MOSF-trypsin catalyzed digestion of myoglobin (2 ng μL^{-1}), and (b) 20 min in-solution digestion of myoglobin (2 ng μL^{-1}).

Table 2 Proteolysis with MOSF-trypsin refrigerated for different time intervals

Days	1	2	3	5	7	9	11	14	18	21
MOWSE score	220	215	221	224	218	223	221	193	166	146
Sequence coverage (%)	89	86	100	99	99	98	88	84	84	81
Peptides matched	15	16	19	22	15	17	19	13	15	15

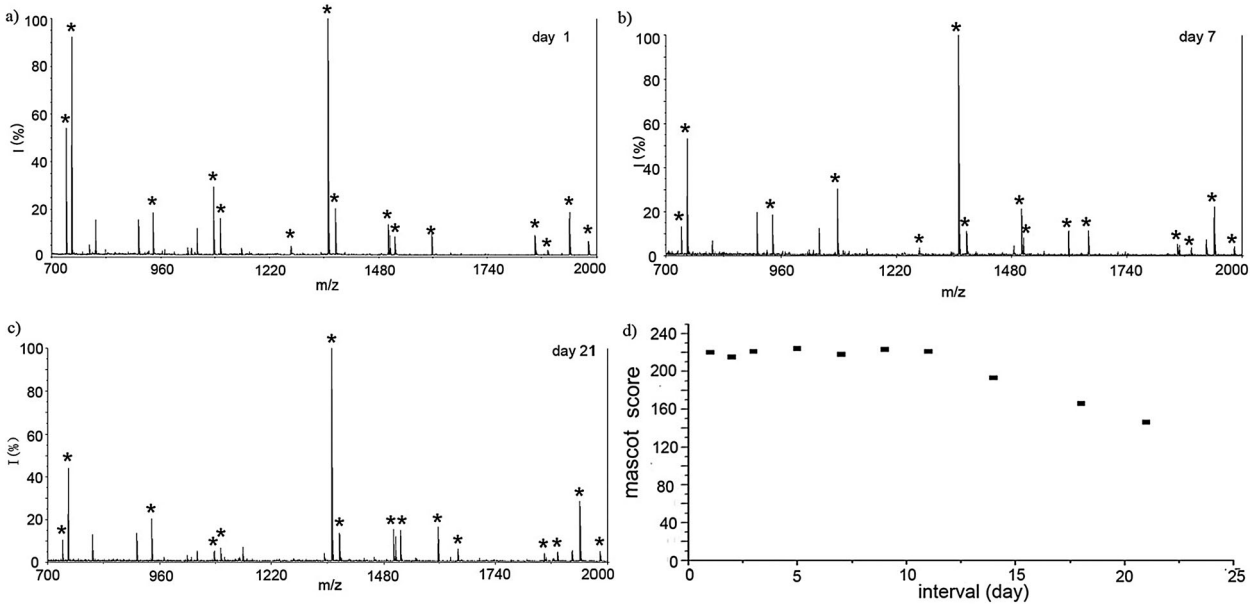

Fig. 4 Mass spectra of myoglobin (20 ng μl^{-1}) proteolysis products catalyzed by trypsin-MOSF reactors stored for (a) 1 day, (b) 7 days and (c) 21 days; (d) MOWSE score of mass spectra plotted against the storage time.

show the activity decline of the immobilized trypsin, Fig. 4d. As shown in Table 2, the digestion with MOSF-trypsin prepared within 2 weeks was always quite efficient with a Mascot searching score above 193. Though the enzyme activity showed some decline after 2 weeks of refrigeration, the myoglobin proteolysis with MOSF-trypsin stored for 3 weeks still showed high efficiency, Fig. 4c. The proteolysis results demonstrate that the immobilized trypsin can be stored for weeks retaining high enzymatic activity. Therefore, the MOSF-trypsin reactor is possible to be produced on a large scale to replace free trypsin in proteome research.

Interference of SDS and urea on the reactor catalyzed digestion

During protein extraction, the pretreatment of complex biological samples and many other biological workflows, surfactants and urea are often used to help protein denaturation and facilitate subsequent steps, which are however hard to be removed completely.[27–29] As impurities, they may severely interfere with the subsequent experimental processes, leading to poor analysis results.[18,30–33] As predicted in SI-2 in the ESI,† since the concentration of positive ions in the pores of silica is ~5 mM, surfactants or urea with high concentrations may enter the macropores of MOSF together with the proteins and further disturb the proteolysis in MOSF. SDS, one of the most used surfactants,[28,29,34,35] was added into the reactor-catalyzed proteolysis system as an interfering factor with three concentration gradients of 0.01%, 0.05% and 0.1% (w/v). As shown in Fig. 5a, b and c, myoglobin could be successfully digested by MOSF-trypsin when the concentrations of SDS were 0.01% and 0.05%. However, when the concentration of

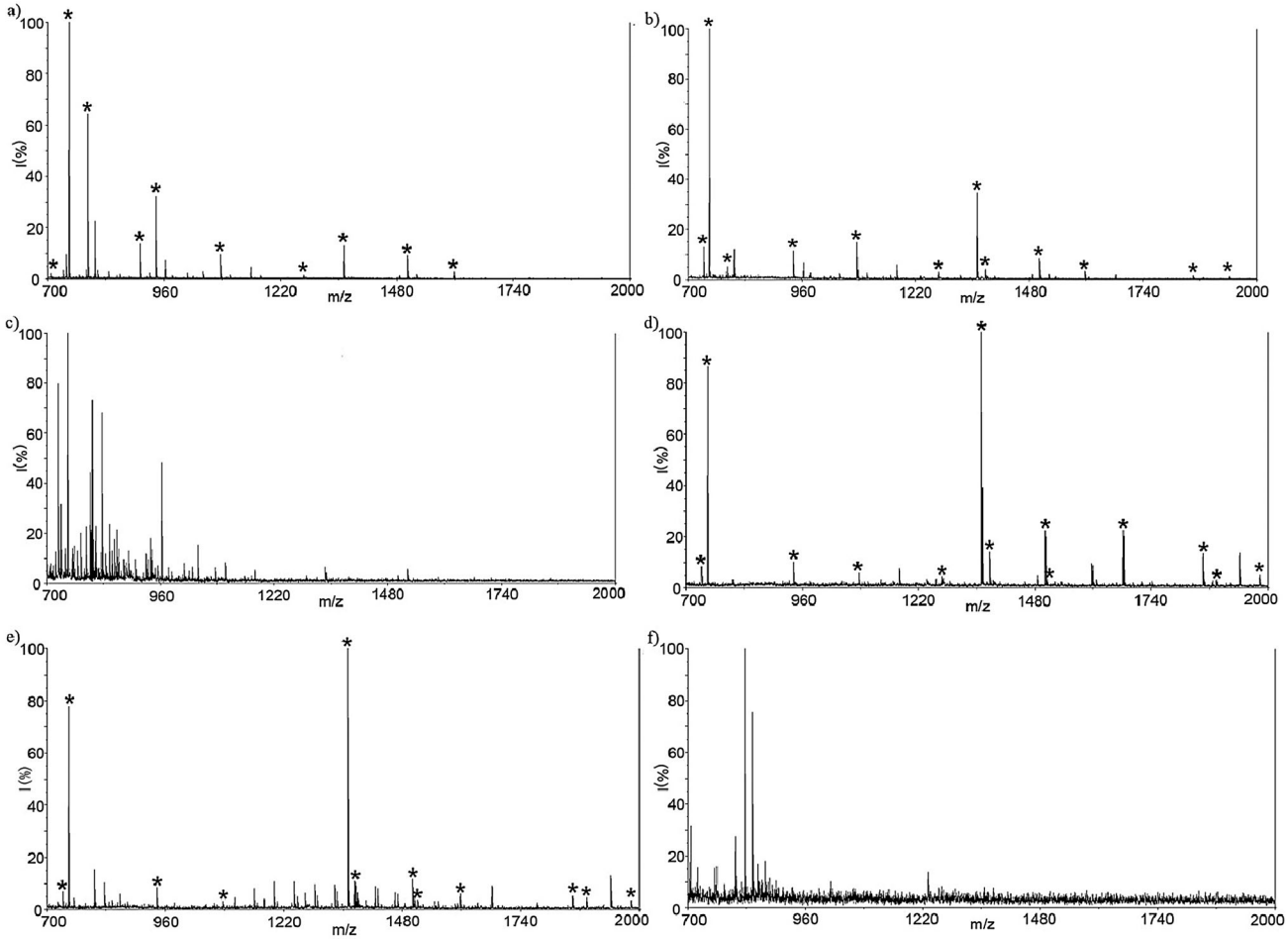

Fig. 5 Mass spectra of proteolysis products from 20 min trypsin-MOSF catalyzed digestion of myoglobin (20 ng μL^{-1}) that is dissolved with (a) 0.01% SDS, (b) 0.05% SDS, (c) 0.1% SDS, (d) 100 mM urea, (e) 200 mM urea and (f) 500 mM urea.

SDS was increased to 0.1%, nothing could be identified by MS from the proteolysis products. Bruening *et al.* also performed a comprehensive study on the effect of SDS on digestion, and suggested that the SDS would affect the proteolysis efficiency of the immobilized trypsin to some extent.[18] It can thereby be concluded that the SDS in the samples should be diluted to a reasonable concentration range before digestion. SDS is usually used at concentrations less than 0.01%.

Urea is also often used to denature proteins. Similar to the SDS test, myoglobin was dissolved in urea solution with concentrations of 100 mM, 200 mM and 500 mM, respectively. As presented in Fig. 5d–f, myoglobin could be successfully digested and identified when the concentrations of urea were 100 mM and 200 mM. When the concentration of urea was increased to 500 mM, the identification failed, which could be a result of either poor proteolysis or bad ionization during MALDI-MS identification. To further investigate the effects of SDS and urea on the digestion efficiency, SDS and urea were removed from the digested samples at the high concentrations of 0.1% w/v SDS and 100 mM urea using SCX Zip Tips. Subsequently, the tryptic peptides were measured by MS, and the identification was found to fail as shown in Fig. SI-2 in the ESI.† The results suggest that the large amount of SDS or urea can decrease the proteolysis efficiency. Therefore, urea or SDS concentrations should be controlled in a certain range to obtain satisfactory results if they could not be removed completely.

Application to complex biological sample analysis

To examine the practical feasibility of the enzymatic silica reactor, MOSF-trypsin was applied to the digestion of a complex biological sample, an HPLC fraction of the proteins extracted from the cytoplasm of mouse liver tissue. After 3 h of digestion with MOSF-trypsin reactors, the tryptic product of 1 μg protein fraction was analyzed by HPLC-ESI-MS-MS. As shown in Fig. 6, 163 proteins were identified from the proteolysis products. In contrast, only 66 proteins were identified from the tryptic peptides generated under the same conditions but with free trypsin (ESI, Table SI-2†). 112 proteins were identified from the proteolysis products by using standard overnight in-solution digestion method, and 82% of these proteins were overlapped by the ones identified using the nanoreactor-based method, indicating only 21 proteins specifically identified, Fig. 6a. As shown in Fig. 6b and c, the molecular weights of the majority of the identified proteins are between 20 000 and 200 000, and their isoelectric points (PI) range from 3.5 to 11 (details in the ESI, Table SI-3†). The number of proteins

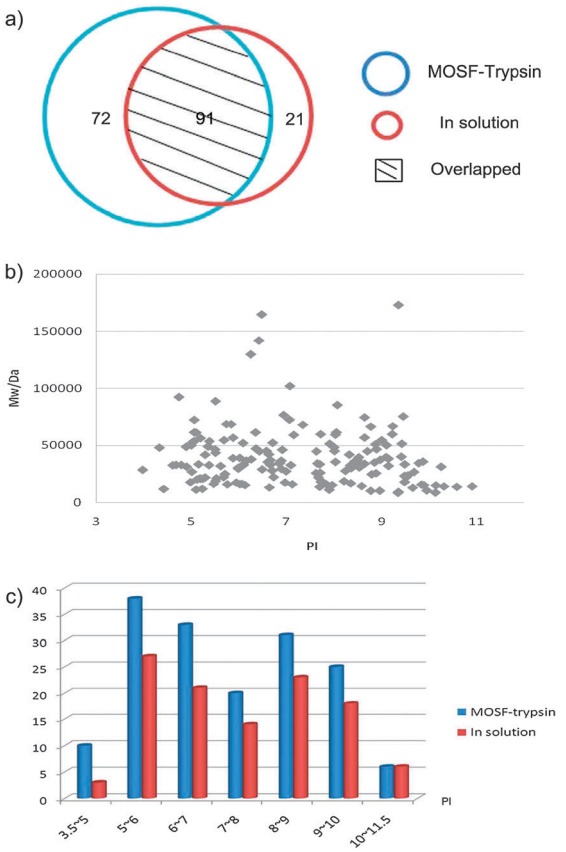

Fig. 6 (a) Comparison of proteins identified from tryptic peptides produced by MOSF-trypsin reactor proteolysis and free trypsin proteolysis of the same biological sample; (b) Plot of PI of the identified proteins as a function of molecular weight; (c) the PI distribution comparison of the proteins identified from tryptic peptides produced by MOSF-trypsin and free trypsin.

identified from MOSF-trypsin-based proteolysis peptides is always more than that from the standard in-solution proteolysis peptides in all PI intervals except the one of 10–11.5. It is quite possible that the peptides with high PI are strongly adsorbed to MOSF. According to our previous work, the zeta potential of MOSF is about −38.8 mV, indicating that its surface is negatively charged when MOSF is dispersed in the proteolysis buffer (pH 8).[22] Therefore MOSF could show electrostatic affinity to the high PI peptides, which are positively charged in the same buffer. With these results, the MOSF-trypsin reactor is demonstrated to be capable of inducing efficient proteolysis of biomolecule substrates with wide ranges of size and PI.

Conclusion

In summary, macroporous silica with large pores is demonstrated both experimentally and theoretically as an efficient host for enzyme immobilization for protein tryptic digestion. A simple, fast and efficient proteolysis approach assisted by MOSF-trypsin reactor is developed and characterized. This silica reactor is stable for weeks and can be easily used to replace the free enzyme during proteolysis while showing an enhanced enzymatic activity. Furthermore, the silica reactor is capable of digesting proteins both at low abundance and in the presence of urea or surfactants. With the application in digesting a biological complex sample, it is demonstrated that the MOSF-trypsin reactor-based protocol could not only be applied to the digestion of standard proteins, but also be an efficient and versatile method for complex biological sample analysis.

Acknowledgements

This work is supported by 973 Program 2007CB714506, NSFC 20925517, and Swiss National Science Foundation 200020–127142. B. H. Liu is grateful to EPFL for a visiting professor fellowship.

Notes and references

1 S. Joshi, A. K. Tiwari, B. Mondal and A. Sharma, *Clin. Chim. Acta*, 2011, **412**, 217–226.
2 J. E. Van Eyk, *Circ. Res.*, 2011, **108**, 490–498.
3 S. Rauser, S. O. Deininger, D. Suckau, H. Hofler and A. Walch, *Expert Rev. Proteomics*, 2010, **7**, 927–941.
4 M. Natesan and R. G. Ulrich, *Int. J. Mol. Sci.*, 2010, **11**, 5166–5184.
5 R. Aebersold and M. Mann, *Nature*, 2003, **422**, 198–207.
6 B. T. Chait, *Science*, 2006, **314**, 65–66.
7 N. L. Kelleher, *Anal. Chem.*, 2004, **76**, 196A–203A.
8 A. Doucette, D. Craft and L. Li, *Anal. Chem.*, 2000, **72**, 3355–3362.
9 D. Craft, A. Doucette and L. Li, *J. Proteome Res.*, 2002, **1**, 537–547.
10 J. Stauber, L. MacAleese, J. Franck, E. Claude, M. Snel, B. K. Kaletas, I. Wiel, M. Wisztorski, I. Fournier and R. M. A. Heeren, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 338–347.
11 X. Y. Yang, Z. Q. Li, B. Liu, A. Klein-Hofmann, G. Tian, Y. F. Feng, Y. Ding, D. S. Su and F. S. Xiao, *Adv. Mater.*, 2006, **18**, 410.
12 J. Deere, E. Magner, J. G. Wall and B. K. Hodnett, *J. Phys. Chem. B*, 2002, **106**, 7340–7347.
13 L. C. Sang and M. O. Coppens, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6689–6698.
14 Y. Liu, H. X. Wang, Q. P. Liu, H. Y. Qu, B. H. Liu and P. Y. Yang, *Lab Chip*, 2010, **10**, 2887–2893.
15 Y. Liu, H. J. Lu, W. Zhong, P. Y. Song, J. L. Kong, P. Y. Yang, H. H. Girault and B. H. Liu, *Anal. Chem.*, 2006, **78**, 801–808.
16 X. J. Xu, X. Y. Wang, Y. Liu, B. H. Liu, H. L. Wu and P. Y. Yang, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 1257–1264.
17 S. Wang, B. G. Wei, P. Y. Yang and G. Chen, *Proteomics*, 2008, **8**, 4637–4641.
18 F. Xu, W. H. Wang, Y. J. Tan and M. L. Bruening, *Anal. Chem.*, 2010, **82**, 10045–10051.
19 H. Y. Bi, L. Qiao, J. M. Busnel, B. H. Liu and H. H. Girault, *J. Proteome Res.*, 2009, **8**, 4685–4692.
20 L. Qiao, Y. Liu, S. P. Hudson, P. Y. Yang, E. Magner and B. H. Liu, *Chem.–Eur. J.*, 2008, **14**, 151–157.
21 H. N. Wang, X. F. Zhou, M. H. Yu, Y. H. Wang, L. Han, J. Zhang, P. Yuan, G. Auchterlonie, J. Zou and C. Z. Yu, *J. Am. Chem. Soc.*, 2006, **128**, 15992–15993.
22 K. Qian, J. J. Wan, L. Qiao, X. D. Huang, J. W. Tang, Y. H. Wang, J. L. Kong, P. Y. Yang, C. Z. Yu and B. H. Liu, *Anal. Chem.*, 2009, **81**, 5749–5756.
23 S. Z. Qiao, C. Z. Yu, W. Xing, Q. H. Hu, H. Djojoputro and G. Q. Lu, *Chem. Mater.*, 2005, **17**, 6172–6176.
24 Y. Qu, L. Moons and F. Vandesande, *J. Chromatogr., Biomed. Appl.*, 1997, **704**, 351–358.
25 Z. Y. Park and D. H. Russell, *Anal. Chem.*, 2000, **72**, 2667–2670.
26 S. Papadopoulos, K. D. Jurgens and G. Gros, *Biophys. J.*, 2000, **79**, 2084–2094.
27 F. M. L. Amado, M. G. Santana-Marques, A. J. Ferrer-Correia and K. B. Tomer, *Anal. Chem.*, 1997, **69**, 1102–1106.
28 Y. X. Zhao, W. Zhang, Y. J. Kho and Y. M. Zhao, *Anal. Chem.*, 2004, **76**, 1817–1823.

29 K. Strupat, M. Karas, F. Hillenkamp, C. Eckerskorn and F. Lottspeich, *Anal. Chem.*, 1994, **66**, 464–470.
30 M. F. Weng and Y. C. Chen, *Rapid Commun. Mass Spectrom.*, 2004, **18**, 1421–1428.
31 A. Gorg, W. Weiss and M. J. Dunn, *Proteomics*, 2004, **4**, 3665–3685.
32 A. H. Brockman, B. S. Dodd and N. Orlando, *Anal. Chem.*, 1997, **69**, 4716–4720.
33 M. Kussmann, E. Nordhoff, H. Rahbek-Nielsen, S. Haebel, M. Rossel-Larsen, L. Jakobsen, J. Gobom, E. Mirgorodskaya, A. Kroll-Kristensen, L. Palm and P. Roepstorff, *J. Mass Spectrom.*, 1997, **32**, 593–601.
34 S. L. Cohen and B. T. Chait, *Anal. Biochem.*, 1997, **247**, 257–267.
35 N. Zhang, A. Doucette and L. Li, *Anal. Chem.*, 2001, **73**, 2968–2975.

